

BBA 95264

ON THE HETEROGENEITY OF THE DEOXYRIBONUCLEIC ACID
ASSOCIATED WITH CRYSTALLINE YEAST CYTOCHROME b_2 J. F. JACKSON, ROGER D. KORNBERG, PAUL BERG, U. L. RAJBHANDARY,
A. STUART, H. G. KHORANA AND ARTHUR KORNBERG*Department of Biochemistry,
Stanford University, Medical Center, Palo Alto, Calif. and
Institute for Enzyme Research, The University of Wisconsin, Madison, Wisc. (U.S.A.)*

(Received January 14th, 1965)

SUMMARY

1. The buoyant density of cytochrome b_2 DNA before and after heating and quick-cooling suggests considerable helicity, and observations on the action of *Escherichia coli* exonucleases I and III on the DNA support this interpretation.

2. Enzymic and chemical investigation of the 3'- and 5'-phosphoryl and 3'- and 5'-hydroxyl end groups of the DNA indicate that these ends are heterogeneous. An average molecular weight range of 20 000–80 000 is estimated from these studies.

3. Cytochrome b_2 DNA serves as a primer for *E. coli* DNA polymerase, thereby permitting a study of overall nearest-neighbor base frequencies. Also, by the use of Mn^{2+} , a single ribonucleotide is incorporated in place of the corresponding deoxyribonucleotide in the product; alkaline hydrolysis of this product releases fragment which can be analyzed for base sequence. Both approaches support the conclusion that the cytochrome b_2 DNA is a heterogeneous population of molecules of the size indicated above or else that to be homogeneous in nucleotide sequence its weight must exceed 100 000.

INTRODUCTION

MORTON and his co-workers obtained cytochrome b_2 , the L(+)-lactate dehydrogenase from baker's yeast, in a crystalline form¹ and identified a DNA component in it². MONTAGUE AND MORTON³ and later MAHLER AND PEREIRA^{4,5} reported this DNA to be about 10 000 in molecular weight and presumably homogeneous in nature. The availability of a DNA of such small size and uniformity of composition would be a boon to studies of enzymic replication and sequence and end group analysis and might prove to have some biological significance as well. For these reasons and with the strong encouragement of the late Professor R. K. MORTON who supplied the material, we undertook an investigation of this DNA. Our results show extensive heterogeneity of the DNA with respect to end groups, chain length and detailed

nucleotide sequences, and confirm the recent report of ARMSTRONG, COATES AND MORTON⁶ that the average molecular weight is considerably greater than originally reported, and perhaps of the order of 100 000.

RESULTS AND DISCUSSION

The cytochrome *b*₂ DNA preparation was generously provided by the late Professor R. K. MORTON and prepared as described by him and MONTAGUE³. Details of the methods used below will not be described inasmuch as they are adequately stated in the cited references.

Helical nature of the DNA

Evidence for a highly ordered conformation of the DNA presented by MAHLER AND PEREIRA⁵ was based on absorbancy changes during thermal treatments. However, the observed lack of interaction with amines or acridine orange and susceptibility to formaldehyde and exonuclease I argued against a typical native, helical structure.

The following three measurements suggest that the DNA has a considerable helical content:

(1) CsCl density gradient centrifugation. The buoyant density measured at pH 7.4 was $1.704 \text{ g} \cdot \text{cm}^{-3}$. After heating the DNA at 100° for 5 min in 0.005 M KCl and quick-cooling, the buoyant density was 1.714. This increase indicates a transition from helical to coil state⁷.

(2) Susceptibility to exonuclease I. This enzyme (also called phosphodiesterase) is specific for single-stranded DNA⁸. When it was used in excess on this DNA sample, 20 % of the DNA-phosphorus was liberated as mononucleotides sensitive to phosphomonoesterase; after heating and quick-cooling the value was 70 %. These results indicate that 20 % of the DNA is single-stranded and accessible to enzymic attack; the failure to achieve complete degradation after heating may be due to partial renaturation or to the presence of 3'-phosphoryl end groups. There is a discrepancy between these results and those of MAHLER AND PEREIRA⁵ showing extensive susceptibility of the unheated DNA to exonuclease I. A possible explanation is that MAHLER AND PEREIRA used conversion to acid-solubility rather than release of nucleoside monophosphates to determine the extent of enzyme action; their assay may also measure oligonucleotides rendered acid-soluble by partial exonuclease action. Another explanation for the discrepancy may be that exonuclease I can contain variable amounts of endonuclease⁹.

(3) Susceptibility to exonuclease III. This enzyme¹⁰ releases 5'-nucleotides from the 3'-hydroxyl ends of double-stranded DNA; the reaction proceeds until about 40 % of the nucleotides are released and the DNA is essentially single-stranded. dAT polymer is more extensively degraded, possibly because even a small fragment can fold back on itself to generate new helical regions. Cytochrome *b*₂ DNA with excess exonuclease III released 80 % of its nucleotide residues, indicating not only that the DNA is helical to begin with, but also that helicity may be restored by the kind of folding process attributed to the dAT polymer.

Number and character of end groups of the DNA

i. *E. coli* alkaline phosphatase. This enzyme will hydrolyze phosphomonoester groups even when they are attached to long polynucleotides. When it is incubated at pH 8 in the absence of Mg^{2+} , several methods fail to detect diesterase (endonuclease) activity. (A stable value of P_1 release is obtained with prolonged incubation and excess enzyme; no single-strand breaks are produced in phage T7 DNA as judged by zonal sedimentation at alkaline pH¹¹.) Several measurements of P_1 release from cytochrome b_2 DNA by the phosphatase gave a value of one mole of P_1 per 70 moles of nucleotide residues. This implies a minimal molecular weight of 20 000 if the average molecule has one terminal phosphomonoester group or 40 000 for two such terminal groups.

2. *DNA phosphatase*. This enzyme which releases P_1 from helical DNA with a phosphomonoester group on the 3'-hydroxyl terminus¹², yielded one mole of P_1 per 140 moles of nucleotide residues in cytochrome b_2 DNA, indicating that a significant number of 3'-hydroxyl ends is esterified with phosphate.

3. *Spleen diesterase*. This enzyme is an exonuclease which releases 3'-nucleotides starting from the 5'-hydroxyl end of a polynucleotide chain; its action is inhibited if this terminus is esterified with phosphate^{13,14}. When it was applied to cytochrome b_2 DNA only 10 % of the nucleotide residues were released; after prior exposure to alkaline phosphatase the value was 60 %. This finding confirms the presence of a significant number of 5'-phosphomonoester termini.

4. *Labeling of phosphomonoester termini with [^{14}C]methyl phosphate*. This method¹⁵ when applied to cytochrome b_2 DNA gave a value indicating one phosphomonoester end group per 200 nucleotide residues (assuming a minimum labeling efficiency of 80 %). When DNA thus labeled was heated and then exposed to exonuclease I, only 70 % of the radioactive label was released in acid-soluble form. The resistance of the remainder of the labeled DNA is concluded to be due to 3'-phosphomonoester end groups. The result of this labeling technique therefore indicates (1) that about 70 % of the total phosphomonoester groups are present as 5'-phosphate end groups in polynucleotide chains which also bear 3'-hydroxyl end groups, and (2) that there are chains which have 3'-phosphate end groups. The latter type of chains may, in addition, carry 5'-phosphate end groups.

5. *Acetylation of end hydroxyl groups and their characterization*. A sample of the cytochrome b_2 DNA was treated with [3H]acetic anhydride¹⁶ and the labeled product, after purification, was degraded, with snake venom phosphodiesterase. After chromatography, 3'-O-acetyldeoxynucleoside 5'-phosphates of adenine, guanine, thymine and cytosine were all identified, demonstrating the presence of free 3'-hydroxyl termini and also their heterogeneity with respect to the terminal nucleoside.

*Cytochrome b_2 DNA as primer for *E. coli* DNA polymerase and the nature of the product*

When this DNA is used as a primer for polymerase, synthesis initially proceeds at about 10 % of the rate observed with dAT polymer. The incorporation of nucleotides into the product (formed up to 40 % net synthesis) shows the same ratio of dTMP to dGMP^{3,5} as that in the primer (2.3). Longer reaction times lead to a rapid increase in this ratio, due to the onset of dAT synthesis¹⁷. When the product isolated after 30 % net synthesis was treated with bacterial alkaline phosphatase,

the release of ^{32}P -labeled inorganic phosphate, where $[\alpha\text{-}^{32}\text{P}]\text{dGTP}$ or $[\alpha\text{-}^{32}\text{P}]\text{dTTP}$ was used as substrate, was 1 mole per 600–1200 total nucleotide residues in the product. This value indicates a molecular weight of 50 000–100 000 for the product, if the average molecule has one phosphomonoester end group, and a distribution of terminal nucleotides which reflects the base composition of the template DNA.

Nearest-neighbor base frequencies¹⁸ in yeast DNA's

Analyses of the cytochrome b_2 DNA and the DNA of baker's yeast show that all the possible dinucleotides are represented and no major discrepancies arise from the sequence frequencies predicted for "random" arrangements (Table I). The rather large deviations between the values for the two bulk yeast DNA's may have considerable significance and require further investigation.

TABLE I

NEAREST-NEIGHBOR SEQUENCES OF YEAST DNA'S

These results were obtained in collaboration with Drs. M. N. SWARTZ and T. A. TRAUTNER.

Nearest-neighbor sequences	Yeast* (1.65)	Cytochrome b_2 (2.14)	Yeast** (2.48)
ApA, TpT	0.097, 0.111	0.115, 0.143	0.122, 0.124
CpA, TpG	0.057, 0.066	0.047, 0.051	0.041, 0.048
GpA, TpC	0.059, 0.058	0.049, 0.054	0.044, 0.045
CpT, ApG	0.057, 0.057	0.054, 0.047	0.040, 0.042
GpT, ApC	0.054, 0.050	0.045, 0.043	0.047, 0.036
GpG, CpC	0.044, 0.038	0.032, 0.036	0.035, 0.032
TpA	0.083	0.113	0.145
ApT	0.100	0.118	0.152
CpG	0.030	0.025	0.024
GpC	0.038	0.029	0.024

* DNA isolated from baker's yeast by Professor R. K. MORTON.

** DNA isolated from Fleischmann's baker's yeast by us. Values in parentheses represent base ratios, A + T/G + C. For further symbols, see ref. 18.

Sequence analysis of the DNA

E. coli DNA polymerase incorporates ribonucleotides in place of deoxyribonucleotides during DNA replication when Mn^{2+} replaces Mg^{2+} as the cation¹⁹, thereby providing the following approach to examining the variety and frequencies of certain nucleotide sequences in the DNA. If $[\alpha\text{-}^{32}\text{P}]\text{CTP}$, dGTP , dATP and dTTP serve as substrates for the replication of a particular DNA, the product formed will contain ribocytidylate at positions normally occupied by deoxycytidylate¹⁹; upon treatment with alkali such a product is cleaved at every cytidylate residue yielding fragments of the general type $(Xp)_n \text{Cp}$ where $n \geq 1$ for $X = \text{dAMP}$, dGMP or dTMP or multiples of these, and where $n = 0$ for $X = \text{dCMP}$. Thus 1 mononucleotide, 3 dinucleotides, 9 trinucleotides and 27 tetranucleotides, etc., are possible products from a completely random sequence of unlimited length (Table II).

The occurrence of ^{32}P in the terminal phosphomonoester residues of the mono-, di- and trinucleotides listed in the first column of Table II shows, that they

TABLE II

POSSIBLE FRAGMENTS PRODUCED UPON ALKALINE CLEAVAGE OF AN ENZYMATICALLY SYNTHESIZED DNA CONTAINING [^{32}P]RIBOCYTIDYLATE

Fragment	Sequences deduced from fragment	
	Synthesized DNA	Template DNA
Cp*	CCC	GGG
Ap* Cp*	CACC	GGTG
Tp* Cp*	CTCC	GGAG
Gp* Cp*	CGCC	GGCG
ApAp* Cp*	CAACC	GGTTG
TpTp* Cp*	CTTCC	GGAAG
GpGp* Cp*	CGGCC	GGCCG
ApGp* Cp*	CAGCC	GGCTG
GpAp* Cp*	CGACC	GGTGG
TpAp* Cp*	CTACC	GGTAG
ApTp* Cp*	CATCC	GGATG
GpTp* Cp*	CGTCC	CGACG
TpGp* Cp*	CTGCC	GGCAG

* The asterisk represents ^{32}P .

arose from tri-, tetra- and pentanucleotide sequences (second column of Table II), respectively, of the newly synthesized polymer. With double-stranded DNA as template, each of the fragments produced during alkaline hydrolysis may be related to a distinct complementary sequence (third column of Table II). Note further that none of the sequences are overlapping so that each fragment corresponds to a unique position in the DNA chain.

For quantitative purposes, the frequency of the sequences shown in columns 2 and 3 can be deduced from the fraction of the monoester ^{32}P relative to the total ^{32}P in each fragment. If, in separate experiments, the same ribonucleotide precursor is used but each of the deoxynucleotide triphosphates is in turn labeled with ^{32}P , the same fragments are produced, but the amount of monoester ^{32}P will be a measure of the frequency with which the labeled deoxynucleotide precursor resides next to the fragments.

This technique has been applied to cytochrome b_2 DNA as primer for DNA polymerase with [α - ^{32}P]CTP, dATP, dGTP and dTTP as substrates. After approximately 15 % net synthesis (the maximum extent to which the synthesis proceeds under these conditions), the product was hydrolyzed with alkali¹⁹, and the digest was chromatographed on DEAE-cellulose according to TOMLINSEN AND TENER²⁰ to yield discrete peaks containing mono-, di-, tri- and tetranucleotides. Each peak was subjected to the two-dimensional mapping procedure of RUSHIZKY AND KNIGHT²¹ and where possible the isomeric trinucleotides were further resolved by chromatography in aqueous ammonium sulfate²¹. The labeled fragments were located by radioautography and eluted, and the ratio of monoesterase-sensitive ^{32}P to total ^{32}P of the fragment was determined using *E. coli* alkaline phosphatase.

With the exception of the isomeric pair ApGpCp and GpApCp, each of the fragments shown in column 1 (Table II) was detected. The ratio of monoesterase-sensitive ^{32}P to stable ^{32}P was always less than 1.0 (usually ranged from 0.10 to

0.30), indicating that each of the sequences represented by the fragment was adjacent to C in only a fraction of the cases: presumably the unlabeled monoesterase-sensitive phosphate was derived from those sequences which were adjacent to the unlabeled dT, dG and dA residues.

Thus, if we assume that cytochrome b_2 DNA is a homogeneous and completely unpaired single-stranded structure, the minimum number of residues serving as template is 55–60. Actually, with the likelihood that any of the 4 nucleotides occurs next to the sequences shown in column 1, the minimum number of residues is 220–240. Moreover, if all of the cytochrome b_2 DNA were part of a Watson–Crick helical structure and copied randomly, one would conclude that there could not be less than 400–500 nucleotide residues per duplex molecule. Even this estimate must be low, since a significant quantity of ^{32}P was found to be associated with tetra-, penta- and higher oligonucleotides. For each tetranucleotide isolated (although these were not identified, there were at least 10 discrete fragments seen on the radioautogram) one would, granting the assumptions discussed above, have to add 24 nucleotides for a single-stranded DNA primer or twice that number for a completely helical primer.

So if cytochrome b_2 DNA contains only 400 residues (120 000 mol. wt.), it cannot be composed of molecules having identical sequences. Alternatively, if it is homogeneous, its size must exceed a molecular weight of $3 \cdot 10^5$.

REFERENCES

- 1 C. A. APPLEBY AND R. K. MORTON, *Biochem. J.*, **71** (1959) 492.
- 2 C. A. APPLEBY AND R. K. MORTON, *Biochem. J.*, **73** (1959) 539.
- 3 M. D. MONTAGUE AND R. K. MORTON, *Nature*, **187** (1960) 916.
- 4 A. S. PEREIRA AND H. R. MAHLER, *J. Mol. Biol.*, **4** (1962) 211.
- 5 H. R. MAHLER AND A. S. PEREIRA, *J. Mol. Biol.*, **5** (1962) 325.
- 6 J. MCD. ARMSTRONG, J. H. COATES AND R. K. MORTON, *Biochem. J.*, **88** (1963) 266.
- 7 N. SUEOKA, J. MARMUR AND P. DOTY, *Nature*, **183** (1959) 1429.
- 8 I. R. LEHMAN, *J. Biol. Chem.*, **235** (1960) 1479.
- 9 I. R. LEHMAN AND A. L. NUSSBAUM, *J. Biol. Chem.*, **239** (1964) 2628.
- 10 C. C. RICHARDSON, I. R. LEHMAN AND A. KORNBERG, *J. Biol. Chem.*, **239** (1964) 251.
- 11 F. W. STUDIER, *J. Mol. Biol.*, **11** (1965) 373.
- 12 C. C. RICHARDSON AND A. KORNBERG, *J. Biol. Chem.*, **239** (1964) 242.
- 13 L. A. HEPPEL AND R. J. HILMOE, in S. P. COLOWICK AND N. O. KAPLAN, *Methods in Enzymology*, Vol. 2, Academic Press, New York, 1955, p. 565.
- 14 W. E. RAZZELL AND H. G. KHORANA, *J. Biol. Chem.*, **236** (1961) 1144.
- 15 U. L. RAJBHANDARY, R. J. YOUNG AND H. G. KHORANA, *J. Biol. Chem.*, **239** (1964) 3875.
- 16 A. STUART AND H. G. KHORANA, *J. Biol. Chem.*, **239** (1964) 3885.
- 17 A. KORNBERG, L. L. BERTSCH, J. F. JACKSON AND H. G. KHORANA, *Proc. Natl. Acad. Sci. U.S.*, **51** (1964) 315.
- 18 J. JOSSE, A. D. KAISER AND A. KORNBERG, *J. Biol. Chem.*, **236** (1961) 864.
- 19 P. BERG, H. FANCHER AND M. CHAMBERLIN, *Symposium on Informational Macromolecules*, Academic Press, New York, 1963, p. 467.
- 20 R. V. TOMLINSEN AND G. M. TENER, *J. Am. Chem. Soc.*, **84** (1962) 2644.
- 21 G. W. RUSHIZKY AND C. A. KNIGHT, *Virology*, **11** (1960) 236.